Refreshing memory traces: thinking of an item improves retrieval from visual working memory

Souza, Alessandra S ; Rerko, Laura ; Oberauer, Klaus

Abstract: This article provides evidence that refreshing, a hypothetical attention-based process operating in working memory (WM), improves the accessibility of visual representations for recall. "Thinking of", one of several concurrently active representations, is assumed to refresh its trace in WM, protecting the representation from being forgotten. The link between refreshing and WM performance, however, has only been tenuously supported by empirical evidence. Here, we controlled which and how often individual items were refreshed in a color reconstruction task by presenting cues prompting participants to think of specific WM items during the retention interval. We show that the frequency with which an item is refreshed improves recall of this item from visual WM. Our study establishes a role of refreshing in recall from visual WM and provides a new method for studying the impact of refreshing on the amount of information we can keep accessible for ongoing cognition.

DOI: https://doi.org/10.1111/nyas.12603

Originally published at:
Souza, Alessandra S; Rerko, Laura; Oberauer, Klaus (2015). Refreshing memory traces: thinking of an item improves retrieval from visual working memory. Annals of the New York Academy of Sciences, 1339(1):20-31.
DOI: https://doi.org/10.1111/nyas.12603

Research Article

Refreshing Memory Traces: Thinking of an Item Improves Retrieval from Visual Working Memory

Alessandra S. Souza [a], Laura Rerko [a], Klaus Oberauer [a]

[a] University of Zurich, Switzerland

Running-head: Refreshing Working Memory Traces

Manuscript Length = 5’156 (including main body of text; but excluding first page, abstract, figure captions, acknowledgments, and references)

Author Note

Alessandra S. Souza, Laura Rerko, and Klaus Oberauer, Department of Psychology, University of Zurich, Switzerland. Correspondence should be addressed to Alessandra S. Souza, Department of Psychology, Cognitive Psychology Unit, University of Zürich, Binzmühlestrasse 14/22, 8050 Zurich, Switzerland. E-mail: a.souza@psychologie.uzh.ch

This article has been accepted for publication in Annals of the New York Academy of Sciences. It has undergone full peer review but has not been through the copyediting, typesetting, pagination, and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as: https://doi.org/10.1111/nyas.12603

Abstract

This article provides evidence that *refreshing*, a hypothetical attention-based process operating in working memory (WM), improves the accessibility of visual representations for recall. "Thinking of" one of several concurrently active representations is assumed to refresh its trace in working memory (WM), protecting it from forgetting. The link between refreshing and WM performance, however, has only been tenuously supported by empirical evidence. Here, we controlled which and how often individual items were refreshed in a color reconstruction task by presenting cues prompting participants to think of specific WM items during the retention interval. We show that the frequency with which an item is refreshed improves recall of this item from visual WM. Our study establishes a role of refreshing on recall from visual WM and provides a new method for studying the impact of refreshing on the amount of information we can keep accessible for ongoing cognition.

Keywords: attention; refreshing; working memory; short-term memory

Introduction

In most quotidian situations we have to maintain information in mind (e.g., our schedule) to efficiently perform an ongoing task (arrange new appointments). Working memory (WM) is the system holding information accessible for on-going cognitive activity. How well information is maintained in WM impacts our ability to perform complex task requiring multiples steps or multiple pieces of information [1–3].

Many theories of WM include the assumption that maintenance of information in WM is accomplished through rehearsal. Whereas verbal information can be maintained by covert articulation (*articulatory rehearsal*) [4,5], it has been claimed that other types of information are maintained through a general, domain-free rehearsal process relying on attention, called *refreshing* [6–9]. Refreshing is defined as the act of turning one's focus of attention to a representation, thereby augmenting its accessibility in WM [10].

Here, we test the conjecture that refreshing improves the accessibility of representations in WM. First, we review the findings used to link refreshing to WM, demonstrating that the evidence for this relationship is, at best, indirect. Next, we present the results of experiments in which we controlled which and how often items were refreshed in visual WM to demonstrate that recall depends on the frequency with which WM contents are refreshed.

Linking Refreshing to WM Performance

Evidence linking refreshing to WM comes primarily from three sources: (1) Manipulations assumed to prevent refreshing; (2) instructions to engage in refreshing; and (3) benefits of guiding attention by cues.

Preventing Refreshing

A commonly held view is that attention is a limited resource. Therefore, in order to refresh memory traces, attention cannot be simultaneously engaged in other processing [6].

Consequently, refreshing would be prevented if responding to an attention-demanding distractor task is required during the retention interval (RI) of a WM task. Increasing the time spent performing such tasks reduces the free time available for refreshing. The ratio of time spent performing distractor tasks over the total RI is referred to as the cognitive load. Numerous experiments have shown that as cognitive load increases, WM performance decreases [11–13]. Although this effect is consistent with a role of refreshing, it is equally consistent with any other hypothetical process that improves memory during a free-time interval. For example, it has been proposed that free time is used to remove distractor information from WM, thereby reducing interference between relevant (e.g., memoranda) and irrelevant information (e.g., distractors). This mechanism has been implemented in a computational model which reproduces the effect of cognitive load on memory that has served as a basis for assuming a role of refreshing in preventing forgetting [14]. Given the fact that refreshing is only one of several possible mechanisms that could be invoked to account for the benefit of increasing free time (i.e., reducing cognitive load), it is important to seek for direct evidence that refreshing contributes to WM performance.

Instructions to Engage in Refreshing

The first and most compelling evidence for refreshing as a distinct mental operation comes from experiments instructing participants to "think of" a memory item: One item is indicated as the “think-of” target, and participants are asked to foreground it in relation to other WM contents [7]. Refreshing one category of items (e.g., scenes over faces) increases brain activity in category-selective areas for the refreshed items and decreases it for non-refreshed items [15,16]. Moreover, refreshing-related brain-activity is distinguishable from activity associated with perceiving the item again [17], and from using articulatory rehearsal to maintain it [10]. In long-term retention tests, refreshed items were better remembered than items presented again for re-reading [18,19]. However, in all behavioral studies the instruction to “think-of” was accompanied by the instruction to recall the refreshed item aloud, so that refreshing was confounded with recall. The improved long-term memory for the refreshed item therefore could be an instance of the testing effect [20]. Moreover, no study so far has investigated the effect of “think-of” instructions on WM tests.

General refreshing instructions ("think of the list items") have also been compared to rehearsal instructions ("repeat the list silently to yourself"); and the findings indicate that refreshing (as opposed to rehearsal) reduces participants’ reliance on phonological representations [21–23]. These studies, however, do not show that refreshing improves WM.

Cueing Benefits

Attention can be guided to one WM item by presenting a retrospective cue during the RI [24]. Access to the retro-cued item is faster and more accurate compared to trials without retro-cues [25–34]. This benefit is consistent with a role of refreshing: focused items are better retrieved from WM than non-focused items. At the same time, the retro-cue benefit is not compelling evidence for a role of refreshing because it is open to many explanations, not all of which are compatible with the notion of refreshing.

Usually, retro-cues indicate with a high validity (70-100%) the to-be tested item; and, by implication, which representations are less relevant (or entirely irrelevant) and could be discarded from WM. The retro-cue benefit therefore could come about because people remove non-cued items from WM to reduce memory load. There is evidence that removal contributes to the retro-cue benefit [34,35]. Other proposed explanations of the retro-cue benefit include the assumption that a cued item is prioritized for comparison to a test stimulus [29] or that the cued item is protected from interference arising from test stimuli [27–29,36,37]. None of these mechanisms would accomplish what refreshing is assumed to do, namely to increase the memory strength of each refreshed item, and to do so potentially for all items in WM.

There is initial evidence that retro-cued items are strengthened in WM: When participants see up to three successive retro-cues (with the last cued item being the relevant for the test), memory is better for items cued twice (sequence of cueing A-B-A) than items cued once (sequence C-B-A) [33]. These findings provide a hint that focusing attention to an item in WM strengthens its representation in a cumulative fashion, so that when the same item is focused again later, further strengthening can build on the effect of previous strengthening, precisely as assumed in WM models assuming refreshing as a maintenance mechanism.

Given that refreshing is only one of the possible contenders to explain cueing benefits, it is important to seek evidence that refreshing improves performance in WM tasks independently of other processes such as removal, prioritization, or protection from interference.

The Present Study

Here, we capitalized on the three sources of evidence reviewed above to develop a method to directly link refreshing and WM performance. First, we provided free time for refreshing. Second, we instructed participants to "think of" specific WM contents; and third, we used (retro-) cues to guide which item was refreshed at which point in time. Unlike traditional retro-cues, our cues did not reliably indicate the item to be tested. This reduces the incentive to remove non-cued items from WM, and hence minimizes the role of removal as an explanation of our findings.

Experiment 1

The goal of this experiment was to test whether items in visual WM that are refreshed more often are recalled better.

Materials and Methods

Participants performed a continuous color reconstruction task [38,39]. Figure 1 illustrates the sequence of events in each trial (further details can be found at the Online Supplementary Materials). Participants were asked to memorize a set of six colors, and to indicate the color of a target item by clicking on a color wheel at the end of a 2.5 s RI. To minimize the contribution of verbal labeling to performance, participants were asked to repeat the sequence *"der-die-das"* aloud throughout the trial.

We inserted a refreshing manipulation during the RI of this task: Four central arrow cues were presented sequentially. Participants were instructed to think of the item each arrow pointed to. Participants were also instructed that the cues did not reliably indicate the item to be tested at the end, but they were urged that thinking of the cued item was part of their main task. We will refer to the presentation of each (retro-)cue as a refreshing step.

There were five possible sequences of cues. Let the letters ABCD represent four diferrent randomly selected memory items: The four arrows could point to four different items (A-B-C-D), they could point to two different items once and to a third item twice (A-B-A-C; A-B-C-B; A-B-C-A), or to two different items twice (A-B-A-B). Note that across two sucessive cues, different items were cued. Through these five sequences we manipulated the frequency with which items were cued as the target of refreshing: 0, 1, or 2 times.

On average, each trial included three items that were not cued (hereafter 0-Refreshing items), two items that were cued once (1-Refreshing items), and one item that was cued twice (2-Refreshing items). In an equal proportion of trials (randomly intermixed), the target of recall was selected to be a 0-, 1-, or 2-Refreshing item. Given these constrains, each 0-Refreshing items had a chance to be tested of 1/9, each 1-Refreshing items had a 1/6 chance, and each 2-Refreshing item had a 1/3 chance to be tested. Therefore, our cues were not totally unpredictive of the item to be tested, but the validity of our cues (maximium 33% for twice-refreshed items) was much lower than previously used in retro-cue studies (~ 67-100% validity [24,40–42]).

Participants were split into two experiments differing only regarding the selection of 1-Refreshing targets. In Experiment 1A (n = 18), 1-Refreshing targets were selected from items refreshed in cue positions 1 or 2; whereas in Experiment 1B (n = 20), 1-Refreshing targets were selected from items refreshed in positions 3 or 4. Participants were not informed about these constrains. This control is critical to test whether participants were able to focus attention on all cued items: if attention was focused only on the last cued items, then 0- and 1-Refreshing targets should not differ from each other in Experiment 1A; whereas in Experiment 1B, 1- and 2-Refreshing targets should not differ from each other. In contrast, if participants used all refreshing cues, and each refreshing step conferred a boost to the refreshed item's accessibility, then a monotonic improvement in recall as a function of refreshing frequency should be observed in both Experiments 1A and 1B.

Data analysis

We computed the distance on the color wheel between the color reported by the participant and the target's true color (ranging from -180º to 180º; hereafter deviation). Our main dependent variable was the absolute value of the deviation which we will call *recall error*. We focused our analyses on testing the effects of two variables on recall error: (1) number of refreshing steps directed to the target; and (2) experiment (1A vs. 1B). We submitted our data to a Bayesian analysis of variance (BANOVA) [43], using the BayesFactor package [44] implemented in R.

The BANOVA computes the strength of the evidence for the presence or absence of an effect by comparing the likelihood of the data given several linear models. Models can be specified including or omitting the main effects and interactions between the variables of interest (fixed effects), while taking into account the effects of nuisance variables (random effects; e.g., participant). The BANOVA computes the likelihood of the data in light of one model – e.g., a Null model assuming that performance is not affected by refreshing (M_0) – in comparison to an alternative model – e.g., a model including a main effect of refreshing frequency (M_1). The BANOVA returns the ratio of the likelihoods of these models (M_1/M_0), which is the Bayes factor (BF). The BF should be interpreted as the multiplicative factor by which our ratio of prior beliefs in the two models should be updated in light of the data. For example, if the $BF_{M1/M0} = 10$, then the data are 10 times more likely under the model assuming an effect of refreshing (M_1) than the model omitting it (M_0), and our ratio of prior beliefs in M_1 over M_0 should be updated by a factor of 10 in favor of M_1.

Results and Discussion

Recall error decreased as the frequency of refreshing steps directed to the target increased (Figure 2a). We entered the recall error in each trial as the predicted variable in the BANOVA. Refreshing frequency and experiment were used as predictor variables. As shown in Table 1, the best BANOVA model included the main effects of refreshing and experiment, but no interaction. The absence of an interaction implies that the effect of experiment reflects differences in overall performance level across samples, but no difference in the effect of refreshing frequency on performance. Moreover, entering refreshing frequency as a linear predictor in a regression model yielded a $BF_{Linear/Null}$ of 1.30×10^{99}, showing a strong linear trend.

We also contrasted individual refreshing levels with each other by replacing the three-level refreshing predictor in the BANOVA with alternative predictors in which pairs of refreshing levels were set to be equal (see Table 1). Pairwise comparisons between the three-level and the alternative predictors are obtained by dividing the BF of the model using the three-level predictor (over the Null) by the BF of the model using the alternative predictor (also over the Null). None of the models using a more constrained predictor yielded a BF larger than the three-level predictor, the three-level predictor being preferred by a factor of at least 11. Finally, we also compared recall of 2-Refreshing targets that were lastly refreshed in positions 3 to those lastly referred in position 4 (Figure 2b): the BF for this comparison favored the Null for the main effect of last refreshing position ($BF_{Null/Last\ Cued}$ = 16.7). Together, these findings rule out the possibility that attention was focused exclusively on the last cued item in the sequence, supporting instead the conclusion that participants followed the “think-of” instruction for all refreshing cues, and each refreshing event improved the accessibility of the refreshed item.

In sum, Experiment 1 showed that recall error decreased monotonically as the number of refreshing steps directed to the target increased from 0 to 1 and from 1 to 2; moreover, the absence of an interaction with experiment shows that this effect was not influenced by where in the cue sequence an item was refreshed.

Experiment 2

What are the consequences of refreshing a subset of the items currently in WM? There are at least three possible scenarios: (1) refreshing one item strengthens the refreshed item above baseline with no change for the non-refreshed items (net benefit hypothesis); (2) the refreshed item is strengthened at the expense of the non-refreshed items (benefit + costs hypothesis); or (3) the refreshed item is maintained at the baseline level, whereas non-refreshed items are weakened (due to decay or interference; net cost hypothesis).

The goal of Experiment 2 was to contrast performance in our refreshing condition to a condition without the insertion of refreshing cues. Our reasoning was that by comparing performance for 0-Refreshing targets against this baseline one could reveal the fate of non-refreshed items in WM, and hence the possible benefits and costs of refreshing. There is one complication in comparing performance across these conditions: It is usually assumed that participants spontaneously refresh items when time permits [6,11,45]. Therefore, the absence of a refreshing manipulation in the baseline does not exclude the possibility that participants are refreshing WM contents in an uncontrolled way. Consequently, even if refreshing only yields a benefit, this comparison would yield the appearance of a cost because the 0-Refreshing items would be refreshed less than the average item in the baseline. To circumvent this difficulty we compared 0-Refreshing targets to recall in a Baseline-short condition in which the RI was equal to the time until presentation of the first refreshing cue in our refreshing condition. This comparison equates the unconstrained refreshing time across conditions. If refreshing strengthens the refreshed items without a cost for non-refreshed items, then 0-Refreshing targets should be recalled as well as targets in the Baseline-short condition, because both targets had the same opportunity for being refreshed.

At the same time, the 0-Refreshing targets from the refreshing condition and the Baseline-short targets differ in RI. We have investigated the effect of RI on memory in the present paradigm and found that it is best described as an effect of temporal distinctiveness [46]. Therefore, we took special care to equate the refreshing condition and the Baseline-short condition with regard to temporal distinctiveness. According to temporal distinctiveness models such as SIMPLE [47], events are discriminated in the psychological dimension of time, which is logarithmically compressed. Therefore, as events recede back in time, their relative distance from each other shrinks, and they become less distinguishable. For example, imagine two events A and B presented 1 s apart from each other. If recall is 1 s after A, the proximity of A and B on the psychological time dimension is the time passed since A over the time passed since B (i.e., ½ = 0.5). If recall is after 5 seconds, this ratio is 5/6 = 0.833. A higher ratio implies higher proximity in psychological time (other things being equal), and hence higher confusability [48]. This interference model predicts that recall gets worse when the RI is increased because the representation of the current memory array becomes less distinct in time from the representation of previous memory arrays. At the same time, recall is predicted to get better when the inter-trial interval (ITI) is increased because it increases the separation of the trial events in time.

In a test of the predictions of temporal distinctiveness theories using the present visual WM task, Souza and Oberauer [46] showed that temporal distinctiveness affects memory performance. Moreover, performance was found to be constant across different levels of RI (1 vs. 3 s) when the relative spacing of the trials (and hence temporal distinctiveness) was held constant (see also Ref. 49). Therefore, we selected the durations of our RIs and ITIs in the refreshing and baseline conditions to match trials in terms of temporal distinctiveness. This control is important to enable comparisons of performance between conditions with different RIs. In particular, if refreshing improves memory with no cost to non-refreshed items, we expect that 0-Refreshing targets (with RI = 3 s) are remembered as well as targets in the Baseline-short condition (with RI = 1 s) when holding temporal distinctiveness constant.

Materials and Method

A new sample of participants (n = 24) completed two sessions. The refreshing session was similar to the one implemented in Experiment 1, with three exceptions. First, 1-Refreshing targets were selected from any cue position. Second, the pre-cue interval was increased to 1 s (total RI = 3 s); and third, the inter-trial interval (ITI) was 7.5 s. The baseline session was divided into the Baseline-short and the Baseline-long conditions. In the Baseline-short condition, RI was 1 s (equal to the pre-cue interval in the refreshing condition) and ITI was 1 s. In the Baseline-long, RI was 3 s, and ITI was 7.5 s. We varied RI and ITI in sync to hold temporal distinctiveness of successive trials constant across all conditions. See Online Supplementary Materials for further details.

Results and Discussion

Recall error in the Baseline-long condition tended to be smaller than in the Baseline-short (Figure 3). However, this difference yielded a $BF_{RI/Null}$ of 2.9, which is weak evidence against the Null hypothesis. Given that in Souza and Oberauer [46], this comparison yielded a $BF_{RI/Null}$ of 0.04 (indicating that the null should be favored by a factor of 25) and that the BF here only slightly favored the alternative hypothesis, we can assume that our control of temporal distinctiveness was reasonably successful. Regarding the Refreshing condition, increasing the number of refreshing steps yielded corresponding decreases in recall error: the BF of the model including refreshing as a predictor exceeded the Null model by 8.1×10^5, and the linear regression yielded a $BF_{Linear/Null} = 6.6 \times 10^{82}$.

Critical to our question is the comparison of the Baseline-short condition to 0-Refreshing targets: the BF of this model was of 0.25, indicating that the Null model should be preferred by a factor of 4. This implies that recall of 0-Refreshing targets was comparable to recall of items from the Baseline-short condition, as expected from the net benefit hypothesis.

Experiment 3

In the previous experiments, the frequency with which an item was refreshed was slightly predictive of the to-be-tested item: The probability of testing 0-, 1-, and 2-Refreshing items in Experiments 1 and 2 was 1/3 for each of these categories. Because on average there was only one 2-Refreshing item, but two 1-Refreshing items, and three 0-Refreshing items in each trial, the chance of the single 2-Refreshing item being tested was larger than that of any individual 1-Refreshing item, which was in turn larger than the chance of any individual 0-Refreshing item. The goal of Experiment 3 was to examine whether totally non-predictive cues can be used to guide the pattern of refreshing.

Materials and Method

Participants (n = 28) completed two sessions with the Refreshing condition. Critically, in this experiment, the target of recall was randomly selected with equal probability among the items in the memory array, so that each item had a 1/6 chance of being tested. Because on average there were three 0-Refreshing items, two 1-Refreshing items, and one 2-Refreshing item per trial, the proportion of tests of the three refreshing categories differed (3/6, 2/6, and 1/6, for the 0-, 1-, or 2-Refreshing categories, respectively). See Online Supplementary Materials for further details.

Results and Discussion

Recall error was slightly reduced by increasing refreshing frequency (see Figure 4), and the BANOVA yielded modest evidence in favor of the main effect of refreshing on recall deviation, $BF_{Ref/Null} = 4.6$. A Bayesian regression with refreshing frequency as a linear predictor yielded strong evidence for the model with a linear effect of refreshing frequency, $BF_{Linear/Null} = 8.4 \times 10^{119}$.

To compare the effect of refreshing across all of our experiments, we plotted the posterior distribution of the regression coefficients obtained in each experiment (see Figure 5). Refreshing led to a smaller benefit in Experiment 3 than the preceding experiments, but the effect was clearly different from zero, and not unambiguously different from the distribution of effects obtained in the previous experiments. Together, these findings show that cueing can be used to motivate participants to selectively refresh the cued items in WM. The extent to which this instruction is effective, however, might depend on the subjective utility of refreshing some WM contents more often than others.

General Discussion

Refreshing has been proposed as a domain-general maintenance mechanism in theories of WM [7,9–11]. So far, a beneficial role of refreshing in WM has never been empirically demonstrated. Our main goal in this study was to examine whether refreshing of individual items improves recall of these items from visual WM. We were successful at guiding the pattern of refreshing during the RI, which allowed us to demonstrate a monotonic improvement in recall performance as refreshing frequency increased. Next, we will discuss potential challenges to our findings, and draw out their implications.

The Speed and the Power of Refreshing

Our study showed only a modest improvement in recall performance due to refreshing: recall improved with each refreshing step, but even after two refreshing steps, performance was far from perfect. There are two reasons why this effect could be an underestimation of the effectiveness of refreshing. First, our cues were presented at a much slower rate (500 ms) than the assumed rate of refreshing (~ 50 ms/item as estimated from set-size slope functions [50]). Therefore, it is possible that participants have sneaked in some refreshing of non-cued items during the RI (at 50 ms/item, refreshing of all items would take only 300 ms!). If participants had refreshed the non-cued items to some extent, that would have worked against the effects we predicted and observed, because it would have diluted the difference in performance between targets cued zero, one, or two times. That said, our finding that 0-Refreshing items were not remembered better than items in the Baseline-short condition (Experiment 2) speaks against the assumption that non-cued items have been refreshed in between refreshing the cued items.

A second consideration to bear in mind is that by using cues to guide refreshing, we are adding a number of processes (e.g., perceiving and decoding the cue, and orienting attention to an item) to the total time of refreshing. Our choice of a rate of 500 ms/item was based on experiments showing that shorter times are less effective at producing retro-cueing benefits [51,52]. A substantial part of the time needed to make use of a retro-cue, or a refreshing cue, might reflect the time to encode and interpret the cue. If this is the case, cue-guided refreshing would be much slower, and hence less efficient, than self-guided refreshing. To conclude, our experiments provide a demonstration that refreshing improves memory, but further work is needed to obtain an estimate of the effectiveness of refreshing when it is used spontaneously.

The Consequences of Refreshing

Experiment 1 suggested that the frequency with which an item is refreshed is more critical for yielding a focusing benefit than the order in which refreshing is directed to an item (from cue 1 to 4). Moreover, non-refreshed items were recalled as well as items from the Baseline-short condition (Experiment 2). These two findings suggest that refreshing yields benefits for the refreshed items without costs for non-refreshed items. This helps ruling out the possibility that our cueing procedure interfered with memory, particularly with memory of non-cued items. More critically, the absence of a cost for non-refreshed items is challenging for models assuming that refreshing serves to restore decayed memory traces, because these models predict that non-refreshed items become weaker over time [6,11,13,45]. Specifically, if non-refreshed items decayed over time whereas refreshed items were protected from decay, we would expect that 2-Refreshing targets should be recalled as well as targets in the Baseline-short condition, whereas 0-Refreshing items should be recalled worse, because they had more time to decay.

One might argue that we are relying on a null finding to claim that refreshing is beneficial. The BF favoring the null model for the comparison of 0-Refreshing targets to Baseline-short targets was only 4 and, hence, not very strong. Therefore, our findings cannot be taken as strongly ruling out the possibility of costs following refreshing (see Ref. [40,41,53]), At the same time, positive, corroborating evidence for a beneficial effect of refreshing over and above maintaining items at their original level of strength can be drawn from comparing the Baseline-short condition to the recall of 2-Refreshing targets. For this comparison the BF = 2406 in favor of the alternative model, indicating that recall of 2-Refreshing targets was better than recall of items in the Baseline-short condition. From this result we can conclude with confidence that refreshing items during the RI improves their accessibility above baseline.

The demonstration of a beneficial effect of refreshing over and above maintenance has an important theoretical implication: The concept of refreshing must be untied from the concept of decay, because refreshing does more than protecting representations from decay. For example, it is possible that refreshing moves items forward in psychological time (i.e., closer to the time of test), so that they gain in temporal distinctiveness relative to non-refreshed items [47]. Alternatively, the bindings between the refreshed item and its retrieval cue might be strengthened by focused attention, thereby facilitating retrieval [8,33,53]. By directly controlling refreshing patterns as we did here, researchers have a chance of contrasting predictions of these accounts, and hence to find out how refreshing works.

Refreshing and Retrieval

In our study we disentangled refreshing from overt recall. However, it is still possible that refreshed items were covertly retrieved. It is difficult to rule out this possibility, but we believe one finding makes this explanation less plausible: Refreshing one item did not impair recall of non-refreshed items This contrasts with what is known from the effects of recalling items from WM: recalling one item reduces the accessibility of not-recalled items [32,34,54]. That said, it is possible that the beneficial effect of refreshing and the testing effect arise from partially overlapping processes preceding retrieval.

Refreshing and the Retro-Cue Benefit

Our demonstration of a refreshing benefit goes beyond previous demonstrations of a retro-cue benefit in three regards. First, we show a beneficial effect of focusing attention on an item in WM even when subsequently other items are focused, demonstrating that the effect of refreshing outlasts the duration for which an item is focused. Second, we show that multiple items can benefit from being refreshed one after another; this benefit cannot be explained by assuming that when one item is focused, all other items are removed from WM; or that only the last focused item is protected from interference from the test situation. Third, we show a beneficial effect of retro-cues with a much lower predictive validity for which item will be tested (at best 33%) than demonstrations of the retro-cue effect (67% or more); and we were able to obtain evidence for a beneficial effect of refreshing even when completely non-predictive cues were used (Experiment 3).

Which Mnemonic Parameter is affected by Refreshing?

We have demonstrated that refreshing reduces the overall error in reporting the color of a target item. Data from the color reconstruction task can be modeled by a mixture model to extract information about theoretical parameters related to the maintenance of items in visual WM, such as the probability that the target was available to be recalled, the precision with which the target's feature is remembered [55], and also the probability of transposition errors (i.e., recall of non-target items) [56]. Given the popularity of this approach, we modelled our data using mixture models (see Online Supplementary Materials). The mixture-model analysis showed that the linear improvement in performance as a function of refreshing is reflected in an increase probability of recalling the target, but not on the precision with which this item is remembered.

Recently the mixture model has been criticized regarding how good a description it provides of the data, and how theoretically informative its parameters are [57–59]. Therefore, given the current debate regarding how to best model responses in the color reconstruction task, we abstain from making any strong claims regarding how refreshing affects underlying memory parameters.

Further Implications

The manipulation of refreshing patterns can be a tool not only to test predictions from WM models, but also for addressing more applied questions. One such promising field relates to the investigation of WM deficits, which could arguably be related to the poor use of refreshing. For example, it has been suggested that young children and old adults do not benefit from refreshing opportunities as much as young adults do [18,53,60–62]; this being one potential reason for their lower performance in WM tasks. Controlling which and how often items are refreshed during the RI can serve to investigate to what extent WM deficits can be attributed to deficits in the effectiveness of refreshing.

Figure Captions

Figure 1. Sequence and timing of events in Experiment 1. ITI = inter-trial interval. In every trial, arrow-cues were used to indicate which item participants should refresh (or, think-of) at which point in time. There was no interval between presentations of successive cues. Note that two successive cues always pointed to different items. Across the four cues, an item could be cued once or twice. Further procedural details can be found in the Online Supplementary Materials.

Figure 2. Results of Experiments 1A and 1B. Panel (a) depicts the mean recall error as a function of the number of refreshing steps (# Refreshings). Panel (b) shows the mean recall error in reporting 2-Refreshing targets that were last refreshed in cue positions 3 and 4. Error bars represent 95% within-subjects confidence intervals [63].

Figure 3. Mean recall error in reporting the target item in the Baseline and Refreshing conditions of Experiment 2. Error bars represent 95% within-subjects confidence intervals [63].

Figure 4. Mean recall error in reporting the target item in Experiment 3. Error bars represent 95% within-subjects confidence intervals [63].

Figure 5. Posterior probability density of the effect of a refreshing step on recall error as estimated from the Bayesian regression analysis of the data of Experiments 1A, 1B, 2, and 3. The 95% credible intervals for the effect in each experiment are shown by the line-bars under each curve. The refreshing effect is negative because it reflects the extent by which recall error is reduced when a refreshing step is directed to an item.

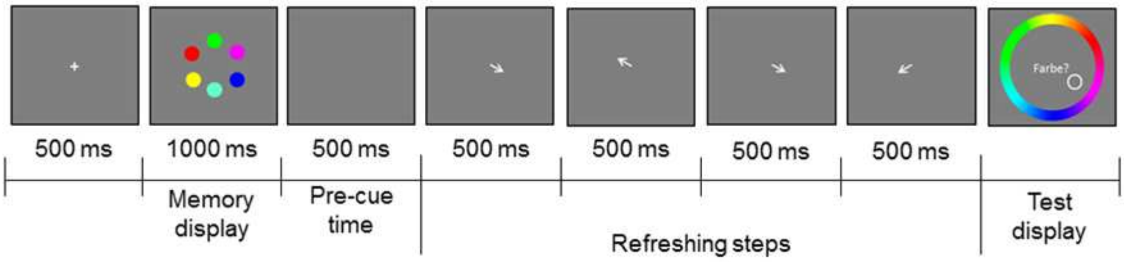

Figure 1. Sequence and timing of events in Experiment 1. ITI = inter-trial interval. In every trial, arrow-cues were used to indicate which item participants should refresh (or, think-of) at which point in time. There was no interval between presentations of successive cues. Note that two successive cues always pointed to different items. Across the four cues, an item could be cued once or twice. Further procedural details can be found in the Online Supplementary Materials.

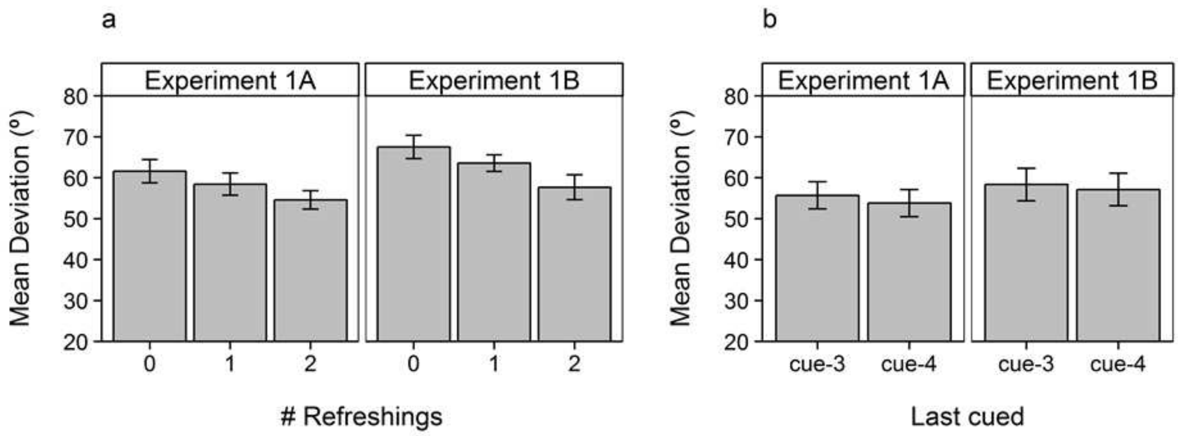

Figure 2. Results of Experiments 1A and 1B. Panel (a) depicts the mean recall error as a function of the number of refreshing steps (# Refreshings). Panel (b) shows the mean recall error in reporting 2-Refreshing targets that were last refreshed in cue positions 3 and 4. Error bars represent 95% within-subjects confidence intervals [63].

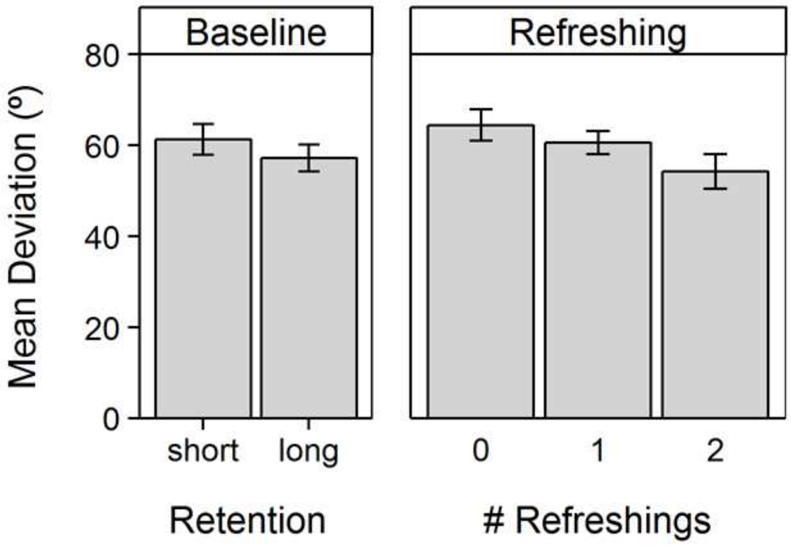

Figure 3. Mean recall error in reporting the target item in the Baseline and Refreshing conditions of Experiment 2. Error bars represent 95% within-subjects confidence intervals [63].

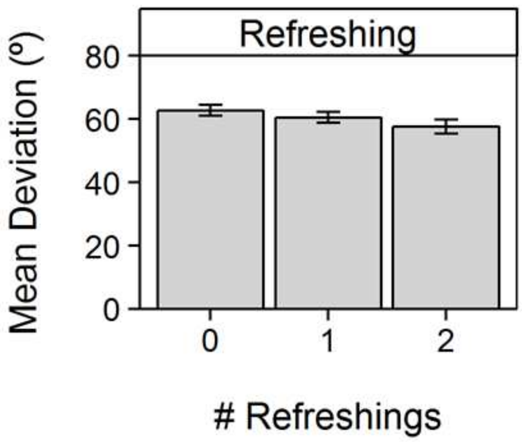

Figure 4. Mean recall error in reporting the target item in Experiment 3. Error bars represent 95% within-subjects confidence intervals [63].

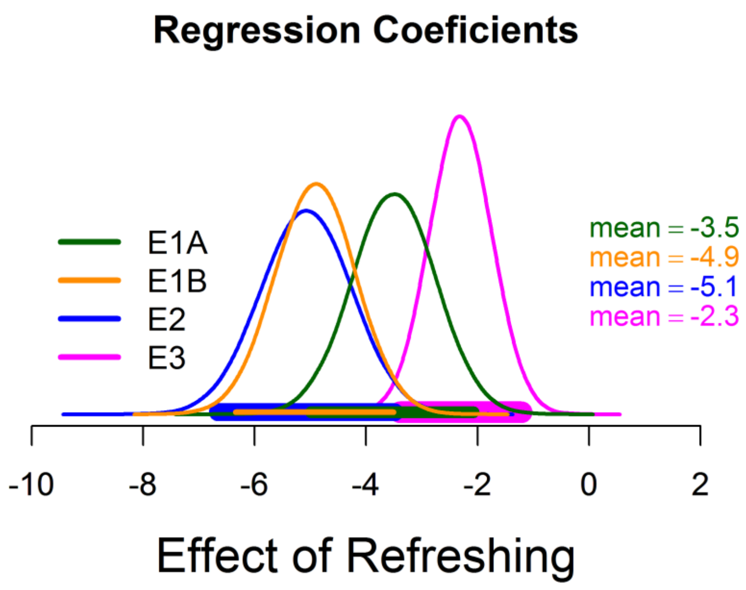

Figure 5. Posterior probability density of the effect of a refreshing step on recall error as estimated from the Bayesian regression analysis of the data of Experiments 1A, 1B, 2, and 3. The 95% credible intervals for the effect in each experiment are shown by the line-bars under each curve. The refreshing effect is negative because it reflects the extent by which recall error is reduced when a refreshing step is directed to an item.

Table 1

Results of the BANOVA with the data of Experiment 1

Refreshing Predictor	Fixed Effects Model			
Model ratio	Ref	Exp	Ref + Exp	Ref + Exp + Ref×Exp
Standard three-level				
[1] 0-Ref ≠ 1-Ref ≠ 2-Ref				
M / Null	1.07×10^{11}	1050	1.15×10^{14}	0.49×10^{12}
Best M / M	1078	1.1×10^{11}	best	239.2
Constrained predictors				
[2] (0-Ref = 1-Ref) ≠ 2-Ref				
M / Null	9.2×10^{9}	1055	0.99×10^{13}	0.60×10^{12}
[1] / [2]	11.63	0.99	11.61	0.81
[3] 0-Ref ≠ (1-Ref = 2-Ref)				
M / Null	4.9×10^{7}	1051	0.53×10^{11}	0.22×10^{10}
[1] / [3]	2083	0.99	2169	222.7
[4] (0-Ref = 2-Ref) ≠ 1-Ref				
M / Null	0.026	1052	26.8	0.7
[1] / [4]	41×10^{11}	0.99	4.29×10^{12}	0.7×10^{12}

Note. M = model; Ref = refreshing predictor [1-4]; Exp = Experiment 1A vs. 1B. All models include participant as a random effect. For tests of each refreshing predictor [1-4], the first row presents the Bayes factor (BF) of each linear model (shown in different columns) over the Null model. For the three-level predictor [1], the second row shows the BF of the best model (highest BF in relation to the Null) over the models presented in different columns. For the constrained predictors [2-4], the second row shows the BF of the model using the three-level predictor [1] over the model using the respective constrained predictor.

Acknowledgment

This research was supported by grants from the Forschungskredit from the University of Zurich (project FK-13-083) to A. S. Souza and from the Swiss National Science Foundation (project 100014_149193) to K. Oberauer. All authors jointly developed the study concept. A. S. Souza programmed the experiments and supervised data collection, which was performed by three research assistants (J. Spengler, M. Thalmann, and S. Colaço). A.S. Souza performed the data analysis and drafted the manuscript. All authors provided critical revisions to the manuscript and approved its final version for submission.

Conflict of Interest

The authors declare no conflicts of interest in publishing this research.

References

1. Kane, M. J., Hambrick, D. Z. & Conway, A. R. A. 2005. Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychol. Bull.* **131:** 66–71
2. Süß, H.-M., Oberauer, K., Wittmann, W. W., Wilhelm, O. & Schulze, R. 2002. Working-memory capacity explains reasoning ability—and a little bit more. *Intelligence* **30:** 261–288
3. Daneman, M. & Carpenter, P. A. 1980. Individual differences in working memory and reading. *J. Verbal Learn. Verbal Behav.* **19:** 450–466
4. Baddeley, A. D., Thomson, N. & Buchanan, M. 1975. Word length and the structure of short-term memory. *J. Verbal Learn. Verbal Behav.* **14:** 575–589
5. Baddeley 1986. Working memory. *Clarendon Press. Univ. Press*
6. Barrouillet, P., Bernardin, S. & Camos, V. 2004. Time constraints and resource sharing in adults' working memory spans. *J. Exp. Psychol. Gen.* **133:** 83–100
7. Johnson, M. K. 1992. MEM: Mechanisms of recollection. *J. Cogn. Neurosci.* **4:** 268–280
8. Loaiza, V. M. & McCabe, D. P. 2012. Temporal–contextual processing in working memory: Evidence from delayed cued recall and delayed free recall tests. *Mem. Cognit.* **40:** 191–203
9. Ricker, T. J. & Cowan, N. 2010. Loss of visual working memory within seconds: The combined use of refreshable and non-refreshable features. *J. Exp. Psychol. Learn. Mem. Cogn.* **36:** 1355–1368
10. Raye, C. L., Johnson, M. K., Mitchell, K. J., Greene, E. J. & Johnson, M. R. 2007. Refreshing: A minimal executive function. *Cortex* **43:** 135–145
11. Barrouillet, P. & Camos, V. 2012. As time goes by temporal constraints in working memory. *Curr. Dir. Psychol. Sci.* **21:** 413–419

12. Vergauwe, E., Barrouillet, P. & Camos, V. 2009. Visual and spatial working memory are not that dissociated after all: A time-based resource-sharing account. *J. Exp. Psychol. Learn. Mem. Cogn.* **35:** 1012–1028
13. Vergauwe, E., Barrouillet, P. & Camos, V. 2010. Do Mental Processes Share a Domain-General Resource? *Psychol. Sci.* **21:** 384–390
14. Oberauer, K., Lewandowsky, S., Farrell, S., Jarrold, C. & Greaves, M. 2012. Modeling working memory: an interference model of complex span. *Psychon. Bull. Rev.* **19:** 779–819
15. Johnson, M. R. & Johnson, M. K. 2008. Top–down enhancement and suppression of activity in category-selective extrastriate cortex from an act of reflective attention. *J. Cogn. Neurosci.* **21:** 2320–2327
16. Johnson, M. R., Mitchell, K. J., Raye, C. L., D'Esposito, M. & Johnson, M. K. 2007. A brief thought can modulate activity in extrastriate visual areas: Top-down effects of refreshing just-seen visual stimuli. *NeuroImage* **37:** 290–299
17. Raye, C. L., Johnson, M. K., Mitchell, K. J., Reeder, J. A. & Greene, E. J. 2002. Neuroimaging a single thought: Dorsolateral PFC activity associated with refreshing just-activated information. *NeuroImage* **15:** 447–453
18. Johnson, M. K., Reeder, J. A., Raye, C. L. & Mitchell, K. J. 2002. Second thoughts versus second looks: An age-related deficit in reflectively refreshing just-activated information. *Psychol. Sci.* **13:** 64–67
19. Raye, C. L., Mitchell, K. J., Reeder, J. A., Greene, E. J. & Johnson, M. K. 2008. Refreshing one of several active representations: Behavioral and functional magnetic resonance imaging differences between young and older adults. *J. Cogn. Neurosci.* **20:** 852–862

20. Roediger, H. L. & Karpicke, J. D. 2006. The power of testing memory: Basic research and implications for educational Practice. *Perspect. Psychol. Sci.* **1:** 181–210
21. Camos, V., Mora, G. & Barrouillet, P. 2013. Phonological similarity effect in complex span task. *Q. J. Exp. Psychol.* **66:** 1927–1950
22. Mora, G. & Camos, V. 2013. Two systems of maintenance in verbal working memory: Evidence from the word length effect. *PLoS ONE* **8:** e70026
23. Camos, V., Mora, G. & Oberauer, K. 2011. Adaptive choice between articulatory rehearsal and attentional refreshing in verbal working memory. *Mem. Cognit.* **39:** 231–244
24. Griffin, I. C. & Nobre, A. C. 2003. Orienting attention to locations in internal representations. *J. Cogn. Neurosci.* **15:** 1176–1194
25. Kuo, B.-C., Stokes, M. G. & Nobre, A. C. 2011. Attention modulates maintenance of representations in visual short-term memory. *J. Cogn. Neurosci.* **24:** 51–60
26. Lepsien, J., Griffin, I. C., Devlin, J. T. & Nobre, A. C. 2005. Directing spatial attention in mental representations: Interactions between attentional orienting and working-memory load. *NeuroImage* **26:** 733–743
27. Makovski, T., Sussman, R. & Jiang, Y. V. 2008. Orienting attention in visual working memory reduces interference from memory probes. *J. Exp. Psychol. Learn. Mem. Cogn.* **34:** 369–380
28. Makovski, T. & Jiang, Y. 2007. Distributing versus focusing attention in visual short-term memory. *Psychon. Bull. Rev.* **14:** 1072–1078
29. Matsukura, M., Luck, S. J. & Vecera, S. P. 2007. Attention effects during visual short-term memory maintenance: Protection or prioritization? *Percept. Psychophys.* **69:** 1422–1434

30. Hollingworth, A. & Maxcey-Richard, A. M. 2013. Selective maintenance in visual working memory does not require sustained visual attention. *J. Exp. Psychol. Hum. Percept. Perform.* **39:** 1047–1058

31. Nobre, A. C., Griffin, I. C. & Rao, A. 2007. Spatial attention can bias search in visual short-term memory. *Front. Hum. Neurosci.* **1:** 4

32. Rerko, L., Souza, A. S. & Oberauer, K. 2014. Retro-cue benefits in working memory without sustained focal attention. *Mem. Cognit.* **42:** 712–728

33. Rerko, L. & Oberauer, K. 2013. Focused, unfocused, and defocused information in working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **39:** 1075–1096

34. Souza, A. S., Rerko, L. & Oberauer, K. 2014. Unloading and reloading working memory: Attending to one item frees capacity. *J. Exp. Psychol. Hum. Percept. Perform.* **40:** 1237–1256

35. Williams, M., Hong, S. W., Kang, M.-S., Carlisle, N. B. & Woodman, G. F. 2013. The benefit of forgetting. *Psychon. Bull. Rev.* **20:** 348–355

36. Sligte, I. G., Scholte, H. S. & Lamme, V. A. F. 2008. Are there multiple visual short-term memory stores? *PLoS ONE* **3:** e1699

37. Pinto, Y., Sligte, I. G., Shapiro, K. L. & Lamme, V. A. F. 2013. Fragile visual short-term memory is an object-based and location-specific store. *Psychon. Bull. Rev.* **20:** 732–739

38. Prinzmetal, W., Amiri, H., Allen, K. & Edwards, T. 1998. Phenomenology of attention: I. Color, location, orientation, and spatial frequency. *J. Exp. Psychol. Hum. Percept. Perform.* **24:** 261–282

39. Wilken, P. & Ma, W. J. 2004. A detection theory account of change detection. *J. Vis.* **4:** 11

40. Astle, D. E., Summerfield, J., Griffin, I. & Nobre, A. C. 2012. Orienting attention to locations in mental representations. *Atten. Percept. Psychophys.* **74:** 146–162

41. Pertzov, Y., Bays, P. M., Joseph, S. & Husain, M. 2013. Rapid forgetting prevented by retrospective attention cues. *J. Exp. Psychol. Hum. Percept. Perform.* **39:** 1224–1231

42. Gözenman, F., Tanoue, R. T., Metoyer, T. & Berryhill, M. E. 2014. Invalid retro-cues can eliminate the retro-cue benefit: Evidence for a hybridized account. *J. Exp. Psychol. Hum. Percept. Perform.* No Pagination Specified doi:10.1037/a0037474

43. Rouder, J. N., Morey, R. D., Speckman, P. L. & Province, J. M. 2012. Default Bayes factors for ANOVA designs. *J. Math. Psychol.* **56:** 356–374

44. Morey, R. D. & Rouder, J. N. 2013. *BayesFactor*. at <http://bayesfactorpcl.r-forge.r-project.org/>

45. Barrouillet, P., Bernardin, S., Portrat, S., Vergauwe, E. & Camos, V. 2007. Time and cognitive load in working memory. *J. Exp. Psychol. Learn. Mem. Cogn.* **33:** 570–585

46. Souza, A. S. & Oberauer, K. 2014. Time-based forgetting in visual working memory reflects temporal distinctiveness, not decay. *Psychon. Bull. Rev.* 1–7 doi:10.3758/s13423-014-0652-z

47. Brown, G. D. A., Neath, I. & Chater, N. 2007. A temporal ratio model of memory. *Psychol. Rev.* **114:** 539–576

48. Lewandowsky, S., Ecker, U. K. H., Farrell, S. & Brown, G. D. A. 2012. Models of cognition and constraints from neuroscience: A case study involving consolidation. *Aust. J. Psychol.* **64:** 37–45

49. Shipstead, Z. & Engle, R. W. 2013. Interference within the focus of attention: Working memory tasks reflect more than temporary maintenance. *J. Exp. Psychol. Learn. Mem. Cogn.* **39:** 277–289

50. Vergauwe, E. & Cowan, N. 2014. A common short-term memory retrieval rate may describe many cognitive procedures. *Front. Hum. Neurosci.* **8:**

51. Souza, A. S., Rerko, L. & Oberauer, K. 2014. Unloading and reloading working memory: Attending to one item frees capacity. *J. Exp. Psychol. Hum. Percept. Perform.* doi:http://dx.doi.org/10.1037/a0036331
52. Tanoue, R. & Berryhill, M. 2012. The mental wormhole: Internal attention shifts without regard for distance. *Atten. Percept. Psychophys.* **74:** 1199–1215
53. Higgins, J. A. & Johnson, M. K. 2009. The consequence of refreshing for access to nonselected items in young and older adults. *Mem. Cognit.* **37:** 164–174
54. Woodman, G. F. & Vecera, S. P. 2011. The cost of accessing an object's feature stored in visual working memory. *Vis. Cogn.* **19:** 1–12
55. Zhang, W. & Luck, S. J. 2008. Discrete fixed-resolution representations in visual working memory. *Nature* **453:** 233–235
56. Bays, P. M., Catalao, R. F. G. & Husain, M. 2009. The precision of visual working memory is set by allocation of a shared resource. *J. Vis.* **9:** 7
57. Van den Berg, R. & Ma, W. J. 2014. 'Plateau'-related summary statistics are uninformative for comparing working memory models. *Atten. Percept. Psychophys.* 1–19 doi:10.3758/s13414-013-0618-7
58. Van den Berg, R., Awh, E. & Ma, W. J. 2014. Factorial comparison of working memory models. *Psychol. Rev.* **121:** 124–149
59. Van den Berg, R., Shin, H., Chou, W.-C., George, R. & Ma, W. J. 2012. Variability in encoding precision accounts for visual short-term memory limitations. *Proc. Natl. Acad. Sci.* **109:** 8780–8785
60. Camos, V. & Barrouillet, P. 2011. Developmental change in working memory strategies: From passive maintenance to active refreshing. *Dev. Psychol.* **47:** 898–904

61. Loaiza, V. M. & McCabe, D. P. 2013. The influence of aging on attentional refreshing and articulatory rehearsal during working memory on later episodic memory performance. *Aging Neuropsychol. Cogn.* **20:** 471–493
62. Loaiza, V. M., Rhodes, M. G. & Anglin, J. 2013. The influence of age-related differences in prior knowledge and attentional refreshing opportunities on episodic memory. *J. Gerontol. B. Psychol. Sci. Soc. Sci.* gbt119 doi:10.1093/geronb/gbt119
63. Cousineau, D. 2005. Confidence intervals in within-subject designs: A simpler solution to Loftus and Masson's method. *Tutor. Quant. Methods Psychol.* **1:** 42–45